J. Clin. Chem. Clin. Biochem.
Vol. 18, 1980, pp. 511–519

A Second External Quality Control Survey (EQCS) for Serum Triiodothyronine (T_3) and Thyroxine (T_4) Assays Using the "Munich Model"[1])

By *W. G. Wood, M. Bauer, K. Horn, I. Marschner, Dagmar van Thiel, Christine Wachter* and *P. C. Scriba*

Laboratorien für Klinische Chemie und Endokrinologie, Medizinische Klinik Innenstadt (Direktor: Prof. Dr. med. E. Buchborn) der Universität München

(Received February 2/March 25, 1980)

Summary: Over 100 laboratories participated in an external quality-control survey (EQCS) for T_3 and T_4, using the Munich Model as conceived in this laboratory ((1–4): *Marschner, I.* et al. (1974), Horm. Met. Res. *6*, 293–296; *Horn, K.* et al. (1976), this J. *14*, 353–360; *Marschner, I.* et al. (1976), this J. *14*, 345–351; *Wood, W. G.* et al. (1980), this J. *18*, 183–192) and carried out over the past 6 years. Twenty lyophilised serum samples, including independant hidden standard curves for T_3 and T_4, were dispatched by post together with a detailed questionnaire and full instructions on reconstitution of the samples. The returned data were processed as previously described ((4): *Wood, W. G.* et al. (1980), this J. *18*, 183–192) and each participant received a full analysis of his own data and a set of histogrammes with which he could visually check his performance against other laboratories. An explanatory letter was sent, which explained the computer print-out, the coding of the kits and contained a constructive report of the participant's performance and helpful advice as to how to improve the assay if this was necessary.

From 110 laboratories returning data for T_3, 86 were fully useable, 22 partly and 2 unuseable because T_3-uptake had been performed instead of T_3-radioimmunoassay (RIA). From 124 laboratories returning T_4 EQCS data, 102 were fully useable and 22 partly useable. In both cases the partly useable sets of data did not contain important items such as the count rates (or absorption values) for each serum, needed for construction and read-off of values from curves I and II. For T_3, intra-assay coefficients of variation (c. v.) of under 5% were seen in 40 cases, and a c. v. of over 15% in 27 cases (range 15–67%). For T_4, 69 participants had a c. v. under 5% and only 7 laboratories had a c. v. above 15% (range 15–41%). The vast majority of participants used commercial kits (T_3 = 90, T_4 = 106).

The results of this EQCS showed a considerable improvement in the five and a half year period since the last EQCS for T_3 and T_4. This is reflected in the intra-assay c. v., especially in the T_3-EQCS where the mean coefficient of variation in 1974 was 15.2% for the three samples and in 1979 for three samples with a similar T_3-content only 31%. Furthermore, some of the kits now on the market show the "ruggedness" desired by *Ekins* ((5): Assay Design and Quality Control: In Radioimmunology 1979 (Ed. Ch. A. Bizollon), Elsevier North Holland, pp. 239–255) that is, kits which give the same results under widely differing conditions.

Ein zweiter Ringversuch für Triiodthyronin (T_3) und Thyroxin (T_4) im Serum

Zusammenfassung: Über 100 Laboratorien nahmen an einem externen Ringversuch für T_3 und T_4 nach dem sogenannten „Münchner Modell", das hier entwickelt ((1–4): *Marschner, I.* et al. (1974), Horm. Met. Res. *6*, 293–296; *Horn, K.* et al. (1976), this J. *14*, 353–360; *Marschner, I.* et al. (1976), this J. *14*, 345–351; *Wood, W. G.* et al. (1980), this J. *18*, 183–192) und in den vergangenen 6 Jahren durchgeführt wurde, teil. Zwanzig lyophilisierte Serumproben einschließlich versteckter Standardkurven für T_3 und T_4 wurden mit der Post, zusammen mit einem detaillierten Fragebogen und ausführlichen Vorschriften für die Proben-Wiederherstellung, versandt. Die zurückgesandten Daten wurden, wie schon beschrieben ((4): *Wood, W. G.* et al. (1980), this J. *18*, 183–192), bearbeitet. Jeder Teilnehmer erhielt eine ausführliche Analyse seiner eigenen Werte und eine Folge von Histogrammen, mit denen er seine Leistung gegenüber anderen Laboratorien überprüfen konnte. Zusätzlich wurde eine Erläuterung zu Computerausdruck und Kit-Verschlüsselung sowie ein konstruktiver Bericht über die jeweilige Teilnehmerleistung versandt und, falls notwendig, ein Rat zur Assay-Verbesserung beigefügt.

[1]) Supported by the Bundesminsterium für Forschung und Technologie (BMFT) D-5300 Bonn 12.

0340-076X/80/0018-0511$2.00

Von den für T_3 erhaltenen Daten aus 110 Laboratorien waren 86 vollständig und 22 teilweise brauchbar. Zwei waren unbrauchbar, da statt eines T_3-RIA ein T_3-Bindungstest durchgeführt wurde. Von den für T_4 erhaltenen Daten aus 124 Laboratorien waren 110 vollständig und 22 teilweise brauchbar.

Beim T_3 sah man einen Intraassay-Variationskoeffizienten (VK) unter 5% in 40 Fällen und über 15% in 27 Fällen (Spannbreite 15–67%). Beim T_4 hatten 63 Teilnehmer einen VK unter 5% und nur 7 einen VK über 15% (Spannbreite 15–41%). Der Großteil der Teilnehmer benutzte kommerzielle Kits (T_3 = 80, T_4 = 106).

Die Ergebnisse dieses Ringversuchs für T_3 und T_4 zeigten seit dem letzten Mal während einer Zeitspanne für fünfeinhalb Jahren einen beträchtlichen Fortschritt. Dies wird deutlich im Intraassay-Variationskoeffizienten, besonders beim T_3-Ringversuch. Bei Proben mit T_3-Konzentration im Normalbereich lag 1974 der Mittelwert der Variationskoeffizienten bei 152%, hingegen betrug er 1979 bei Proben mit ähnlicher Konzentration nur 31%. Ferner zeigen einige der Kits, die jetzt auf dem Markt sind, die „Robustheit", die *Ekins* ((5): Assay Design and Quality Control: In Radioimmunology 1979 (Ed. Ch. A. Bizollon), Elsevier North Holland, pp. 239–255) forderte, d. h. die selben Ergebnisse trotz Durchführung unter verschiedenen Bedingungen.

Introduction

In 1974 thirty members of the German Society for Endocrinology, Thyroid Section, were invited to take part in an EQCS for T_3 and T_4 in serum. The results of the survey were published in detail in 1976 (2). The results showed that the state of these assays was far from ideal. Five years later, a similar study has been organised in which over 100 participants from all types of laboratory have taken part. The method for organising and working out the data received has recently been published (4). The major change in the EQCS-form is that the samples were for the first time, in a survey carried out from this laboratory, dispatched in a lyophilised form. The methods used for the measurement of the thyroid hormones in this EQCS included RIA, EMIT (enzyme multiplied immunoassay technique), ELISA (enzyme linked immunosorbent assay) and CPBA (competitive protein binding assay) techniques.

Materials and Methods

Hidden standard curves were set up in T_3- und T_4-free human serum (0-serum) kindly supplied by commercial kit producers (Henning, Berlin; Amersham-Buchler, Braunschweig). T_3, T_4 and reverse T_3 (rT_3) were obtained from Henning, Berlin, in which a purity of > 99.9% was given. To obtain full solution of these compounds, they were first dissolved in dilute NaOH before being diluted and added to the 0-serum.

Standard curves for T_3 ranged from 0–11 nmol · liter^{-1} and for T_4 from 0–300 nmol · liter^{-1} each in 6 steps. Apart from the 0-standard, different human serum samples were used for the two standard curves. Serum pools were collected from patients with elevated thyroxine binding globulin (TBG) levels as well as from patients with normal and reduced levels. Elevated TBG sera came from women under oestrogen therapy, and reduced TBG serum from a patient with haemochromatosis treated by regular venipunction. rT_3 was added to one serum in a concentration which is often seen in fasted patients (6), and the effect of ion-exchange removal of T_3 and T_4 from a pool serum from normal volunteers was tested in yet another serum. The intra-assay c. v. was tested using a pool from normal volunteers in three different vials.

All sera were filtered under 5 Bar N_2 through an asbestos filter to remove debris and bacteria.

Table 1 shows the composition of the 20 serum samples sent to each participant. All sera were tested for TBG-levels, total protein and electrophoretic mobility after treatment. The capability of the serum to bind radioactive T_4 was also checked as an indirect measurement of the TBG-binding capacity. Except in the specific cases, TBG-levels were normal, as were serum total protein and electrophoresis.

Tab. 1. Composition of the 20 serum samples for the T_3- and T_4-EQCS.

Serum No.	Contents
1	T_3 and T_4-free serum – zero standard for hidden standard curve (TBG 16 mg · liter^{-1}, total protein 69 g · liter^{-1})
2	Commercial quality control serum – Gödecke Validate-N
3	Poolserum extracted with anion-exchange resin, with additional rT_3 end concentration 15 nmol · liter^{-1} (for TBG and total protein see serum no. 7)
4	Poolserum – in normal range, for intra-assay c. v. (TBG – 19 mg · liter^{-1}, total protein – 69 g · liter^{-1})
5	Poolserum for intra-assay c. v. – see serum no. 4
6	Hidden standard curve – T_3 – 5 nmol · liter^{-1}
7	Poolserum – see sample 3 for details, without added rT_3 (TBG – 16 mg · liter^{-1}, total protein 60 g · liter^{-1})
8	Commercial quality control serum – Gödecke Validate-A
9	Hidden standard curve – T_4 300 nmol · liter^{-1}
10	Hidden standard curve – T_3 11 nmol · liter^{-1}
11	Hidden standard curve – T_4 120 nmol · liter^{-1}
12	Serum with reduced TBG-content (TBG – 9 mg · liter^{-1} Total protein – 66 g · liter^{-1})
13	Hidden standard curve – T_3 – 1 nmol · liter^{-1}
14	Hidden standard curve – T_3 – 2 nmol · liter^{-1}
15	Hidden standard curve – T_4 – 60 nmol · liter^{-1}
16	Hidden standard curve – T_4 – 30 nmol · liter^{-1}
17	Poolserum for intra-assay c. v. – see serum no. 4
18	Hidden standard curve – T_3 – 0.5 nmol · liter^{-1}
19	Poolserum from women on oestrogen therapy (TBG – 54 mg · liter^{-1}, total protein – 77 g · liter^{-1})
20	Hidden standard curve – T_4 – 10 nmol · liter^{-1}

All serum samples for the hidden standard curve have the same TBG- and total protein content as serum no. 1.

TBG and Total Protein were not measured in serum nos. 2 and 8.

Returned data was transferred to punched cards (4) and processed in a computer (Siemens 404/3).

Table 2 shows the list of commercial kits used in this survey and the code numbers allotted to them, together with the major features.

Tab. 2. List of Kits used in the T_3- and T_4-EQCS.

Code Number	Kit manufacturer and assay method
None	Laboratory own method irrespective of assay method
1 (51)	Clinical Assays – RIA – Coated Tubes
2 (52)	Diagnostic Products Corporation – RIA – Double antibody with polyethylene glycol (PEG) separation
3	Amersham – RIA – PEG separation
4 (54, 30)	Corning – RIA – solid phase first antibody
5	Byk-Mallinckrodt – RIA – Ion exchange separation
6	Merck (Syva) – EMIT – homogeneous system
10	Beckman – RIA – Double antibody separation
11 (61)	Henning – RIA – Double antibody separation
12	Abbott – RIA – PEG separation
14 (64)	Behring – RIA – PEG separation
15	Becton-Dickinson – RIA – Coated Tubes
16	Pharmacia – RIA – Solid phase first antibody
17	Serono (Biodata) – RIA – PEG separation
18	BioRad – RIA – solid phase first antibody
20	Squibb Clasp – RIA – Coated Tubes
21	Byk-Mallinckrodt SPAC – RIA – Coated Tubes
22	Micromedic (Autopak) – RIA – Coated Tubes
23	Biomerieux – RIA – Double antibody with PEG separation
24	Boehringer – RIA – Coated Tubes
25	Boehringer – ELISA – Coated Tubes
26 (33)	Amersham (Thyopac) – CPBA – Column separation
27 (77)	Ames (Seralute) – RIA – Column separation
28	Henning (RIAcid) – RIA – Preincubated 1st and 2nd antibody
29	IMO – RIA – Coated Tubes
31	Henning – RIA – Charcoal separation
32	Union Carbide (Centria) – RIA – Double antibody separation

The numbers in brackets for kits 30 and 33 are most probably the same as kits 4 and 26 respectively. Kits numbered above 51 are modifications to the methods in the kits. e. g. Kit 51 is the same as kit 1, but the user has modified the method in some way or other.

Results

Participants – According to Laboratory Type and Size, together with Returned Data

Table 3 shows the returned data from participants and whether they could be fully or only partly evaluated. The table is split up into laboratory type and size.

The figures in the square brackets indicate the percentage of all laboratories, those in the round brackets the percentage of the group in question. The table shows that around 80% of all participants could fill out the questionnaires and send back their data in the required form. This ability was seen to the same extent in both general and radiological laboratories. In the cases of partly useable results, data failed, thus preventing the construction of curves I and II, and restricting the data returned to the participant.

Performance (Intra-assay c. v.) – According to Laboratory Type and Size

Table 4 shows the intra-assay c. v. in terms of laboratory type and size. The figures in square brackets represent the percentage within each group and the round brackets have the same meaning as in table 3 above. In the T_3-EQCS, 36% of the participants had an intra-assay c. v. of under 5% and 25% an intra-assay c.v. of over 15%. The same data for the T_4-EQCS were 56% and 6% respectively. A difference in performance between general laboratories performing radioimmunoassays and those specialising in nuclear medicine was not to be seen.

Performance – Own Methods and Commercial Kits

Table 5 shows the performance in the T_3-EQCS in terms of all participants, those using their own method, and for the kits with 5 or more sets of returned data (please refer to table 2 for the kit code).

As the amount of data here is enormous, two things have been done, namely the sera have been organised into ascending concentration order expressed by the group mean, and the simplification of data into table 6 which shows the mean c. v. for the ranges shown. Table 6 shows a "precision-profile" in simplified form. Tables 7 and 8 show the same procedure for the T_4-EQCS.

Outliers in both T_3- and T_4-EQCS are underlined in both tables 5 and 7 where accuracy is questionable. The 3 samples for the intra-assay c. v. (samples 4, 5 and 17) are outlined to allow an easy comparison to be made.

Crossreaction of T_3, T_4 and rT_3

From the results in table 5 it can be seen that for the highest T_4-standard (sample 9—300 nmol · liter^{-1}) only one kit (kit 4) showed a crossreaction giving a "T_3" value of 1.71 nmol · liter^{-1}. In the normal range of T_4 (samples 15 and 11—60 and 120 nmol · liter^{-1}) cross-reaction was not high enough to falsify results. Again kit 4 had the highest cross-reactivity. No T_3-kit showed demonstrable crossreaction with rT_3 (samples 3 and 7).

In the T_4-EQCS crossreaction of T_3 played no role due to the concentration difference. The same was true for rT_3, although almost all antibodies crossreacted with rT_3 (samples 3 and 7, table 7).

Regression Data Analysis

Table 9 shows the regression data for both T_3- and T_4-EQCS. In the first instance for the comparison of data read off the participant's own standard curve using his own method for working out the results (own method) and the same data after standardisation using a spline function (7) (curve I). Any deviation from the ideal value of the regression line $y = a + bx$ where $a = 0$ and $b = 1$, stems from a difference in the working out

Tab. 3. Analysis of returned data – useability for computer processing, according to laboratory size.

Laboratory type	T_3-EQCS			T_4-EQCS		
	Data fully useable	Data partly useable	Data unuseable	Data fully useable	Data partly useable	Data unuseable
General						
Small Private	21	9	0	28	9	0
Large Private	8	0	0	12	0	0
University	8	1	0	8	1	0
State Hospitals	8	2	2	10	2	0
Kit Producers	13	2	0	16	2	0
Others	2	0	0	2	0	0
Radiology/Nuclear Medicine						
Small Private	2	2	0	2	2	0
Large Private	4	0	0	2	2	0
University	11	2	0	11	2	0
State Hospitals	9	4	0	11	2	0
Totals						
All [%]	86 [78]	22 [20]	2 [2]	102 [82]	22 [18]	0
General (%)	60 (79)	14 (18)	2 (3)	76 (84)	14 (16)	0
Radiology (%)	26 (76)	8 (24)	0	26 (76)	8 (24)	0

The square brackets represent the percentage of all laboratories, whereas the round brackets represent the percentage of the two groups in question.

Tab. 4. Analysis of returned data – Intra-assay c. v. – according to laboratory size and type.

Laboratory type	Intra-assay c. v. – T_3-EQCS				Intra-assay c. v. – T_4-EQCS			
General	< 5%	5–10%	10–15%	> 15%	< 5%	5–10%	10–15%	> 15%
Small Private	10 [33]	8 [27]	4 [13]	8 [27]	24 [65]	10 [27]	2 [5]	1 [3]
Large Private	5 [62]	1 [12]	1 [12]	1 [12]	5 [42]	6 [50]	0	1 [8]
University	3 [33]	1 [11]	2 [22]	3 [33]	4 [44]	3 [33]	2 [22]	0
State Hospitals	2 [17]	5 [42]	1 [7]	2 [17]	5 [42]	2 [17]	4 [33]	1 [8]
Kit Producers	8 [53]	3 [20]	1 [7]	3 [20]	12 [67]	5 [28]	1 [5]	0
Others	1 [50]	1 [50]	0	0	1 [50]	1 [50]	0	0
Radiology/Nuclear Medicine								
Small Private	2 [50]	1 [25]	1 [25]	0	1 [25]	3 [75]	0	0
Large Private	0	0	0	4 [100]	1 [33]	1 [33]	0	1 [33]
University	5 [38]	2 [15]	1 [8]	5 [38]	7 [54]	2 [15]	2 [15]	2 [15]
State Hospitals	4 [31]	8 [62]	0	1 [8]	9 [69]	2 [15]	1 [8]	1 [8]
Totals								
All [%]	40 [36]	30 [27]	11 [10]	27 [25]	69 [56]	35 [28]	12 [10]	7 [6]
General (%)	29 (39)	19 (26)	9 (12)	17 (23)	51 (57)	27 (30)	9 (10)	3 (3)
Radiology (%)	11 (33)	11 (33)	2 (6)	10 (29)	18 (55)	8 (24)	3 (9)	4 (12)

The square brackets represent the percentage of each group of laboratories – these figures when read horizontally equal 100%. The round brackets represent the percentage of the two main groups – *General* and *Radiology*, and can be used in comparing performance directly.

method or data transformation process. The second instance shows a comparison of standardised data from the participant's standard curve (curve I) compared with the values read off the hidden standard curve (curve II). Differences occurring through methodology are seen in cases where a differs widely from 0 and b deviates widely from 1. In such cases, where the correlation coefficient remains highly significant, a systematic error is indicated. Where the value of the correlation coefficient (r) falls, random errors come into play.

The boundaries for a, b and r in table 9 were chosen arbitrarily. The results seen in this table are probably the most important in the whole survey, and demonstrate that a good correlation exists between own method of data handling and the spline function used to standardise the data transformation, at least in the majority of laboratories (T_3 = 88%, T_4 = 89%). In the case of both surveys most of the participants had a regression line slope within 10% of the expected value of 1.00 (T_3 = 81%, T_4 = 86%) which in a similar propor-

Tab. 5. Means (nmol · liter^{-1}) and c. v. (%) of methods allowing statistical analysis (n ≥ 5), T_3-EQCS. Sera arranged in ascending concentration of the mean values from all participants. Outliers are underlined, and the 3 samples for the intra-assay c. v. are framed (sera 4, 5 and 17).

Serum No.	All n = 90		Own method n = 19		Kit 2 n = 10		Kit 3 n = 5		Kit 4 n = 5		Kit 10 n = 9		Kit 11 n = 8		Kit 12 n = 5		Kit 14 n = 10		Hidden standard curve Added T_3 (nmol·liter^{-1})
	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	
20	0.27	59.2	0.18	39.0	0.21	38.1	0.17	11.8	0.17	(n = 1)	0.50	8.0	0.68	(n = 2)	0.18	27.7	0.23	26.1	
1	0.31	58.1	0.22	27.2	0.19	26.3	0.17	5.9	0.11	(n = 1)	0.62	9.7	0.46	(n = 1)	0.17	23.5	0.22	28.7	0
16	0.31	54.8	0.25	40.0	0.22	13.6	0.18	11.1	0.18	44.4	0.62	24.2	0.42	116	0.24	12.5	0.29	31.0	
15	0.38	50.0	0.33	39.4	0.29	17.1	0.21	19.0	0.33	27.3	0.68	14.7	0.43	114	0.36	13.9	0.32	21.8	
11	0.52	44.2	0.53	58.4	0.48	37.5	0.31	38.7	0.79	26.5	0.82	8.5	0.40	52.5	0.58	5.2	0.41	34.2	
18	0.55	45.4	0.53	49.1	0.56	16.1	0.49	57.1	0.37	62.2	1.03	8.7	0.30	43.3	0.52	7.6	0.56	34.0	0.5
9	0.93	37.6	0.75	52.0	0.75	12.0	1.11	17.1	1.78	36.5	1.16	13.7	0.74	27.3	1.09	5.5	0.71	28.1	
13	1.02	32.4	0.87	41.3	0.95	7.4	1.31	16.8	0.67	41.8	1.61	6.2	0.75	24.0	1.01	7.0	1.19	11.8	1.0
17	1.01	31.7	0.86	36.0	0.83	10.8	1.31	11.4	1.22	50.8	1.58	5.7	0.92	32.2	1.23	4.7	0.99	11.0	
4	1.03	30.1	0.86	29.1	0.86	7.0	1.61	45.9	1.18	29.7	1.51	10.6	0.83	22.9	1.22	2.5	0.97	17.5	
5	1.03	29.1	0.91	36.3	0.89	11.2	1.27	14.2	0.93	20.4	1.51	8.6	0.91	30.8	1.16	5.2	0.91	25.3	
3	1.10	29.1	1.06	29.2	1.08	5.5	1.12	15.2	0.95	25.3	2.31	6.1	1.08	39.8	1.19	16.8	0.89	30.3	
7	1.13	25.6	1.04	24.0	1.14	7.9	1.41	47.0	0.72	45.8	2.17	11.0	1.16	14.7	1.17	12.8	0.99	15.2	
8	1.47	25.8	1.29	31.0	1.47	12.9	1.68	12.5	1.62	19.2	1.80	4.4	1.25	20.0	1.95	10.3	1.13	14.2	
2	1.54	23.4	1.32	23.2	1.49	1.6	1.87	8.5	1.36	19.8	1.99	8.0	1.32	5.3	1.71	8.7	1.41	23.4	
12	1.69	19.5	1.62	16.7	1.71	10.1	1.95	16.4	1.60	21.2	2.31	8.6	1.53	18.9	1.97	9.1	1.52	11.8	
14	1.99	18.6	1.83	24.6	1.87	3.2	2.48	15.7	1.74	15.5	2.69	3.3	1.77	10.2	2.02	9.4	2.13	13.5	2.0
19	2.80	17.4	2.76	20.6	2.49	4.8	3.33	14.7	3.00	30.3	3.36	2.3	2.75	18.2	2.94	6.8	2.73	15.4	
6	5.39	14.5	5.20	14.2	4.97	5.2	6.77	9.8	4.89	6.3	6.60	8.5	5.96	7.9	5.24	9.4	5.09	21.8	5.0
10	11.2	11.1	11.6	9.1	10.9	7.2	13.0	(n = 1)	10.7	3.8	11.6	(n = 2)	10.7	26.9	11.2	7.4	7.93	(n = 1)	11.0

Tab. 6. Mean coefficients of variation – Precision profile – T_3-EQCS.

Range (nmol · liter^{-1})	All n = 90	Own method n = 19	Kit 2 n = 10	Kit 3 n = 5	Kit 4 n = 5	Kit 10 n = 9	Kit 11 n = 8	Kit 12 n = 5	Kit 14 n = 10
mean c.v.									
(0–11)	32.9	32.0	12.8	20.5	29.2	9.02	34.7	10.3	21.8
0–1	49.9	40.7	17.9	23.9	36.7	13.0	51.4	15.7	25.3
1–2	26.5	24.8	6.87	20.5	27.5	8.24	18.2	8.26	15.3
> 2	14.3	14.6	5.73	13.4	13.5	6.63	17.7	8.25	16.9

The table summarises the data from table 5 in different regions of the standard curve. It must be emphasized that this table shows the precision but not the accuracy of a method.

Tab. 7. Means (nmol · liter^{-1}) and c.v. (%) of methods allowing statistical analysis (n ≥ 5) T_4-EQCS. Sera arranged in ascending concentration of the mean values from all participants. Outliers are underlined and the 3 sera for the intra-assay c.v. are framed.

Serum No.	All n = 103		Own method n = 18		Kit 1 n = 6		Kit 2 n = 10		Kit 3 n = 5		Kit 4 n = 8		Kit 6 n = 9		Kit 11 n = 8		Kit 12 n = 7		Hidden standard curve Added T_4 (nmol·liter^{-1})
	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	mean	c.v.	
14	8.49	60.1	7.77	55.0	7.67	44.1	5.75	33.0	6.29	37.3	14.0	34.0	18.1	42.2	3.46	77.5	6.50	50.6	
18	8.79	68.5	9.44	59.1	6.91	30.2	5.57	31.6	7.52	26.4	13.3	56.9	24.0	50.2	4.05	55.8	5.89	67.7	
13	9.07	66.6	8.48	57.4	8.27	52.8	6.06	13.9	6.75	31.8	15.2	37.5	19.5	37.0	5.24	24.9	5.46	33.9	
10	9.18	81.9	8.05	54.7	5.73	74.7	4.43	42.7	5.51	45.9	13.0	8.1	19.5	12.6	3.99	101	5.84	34.1	
6	9.55	70.3	10.0	46.5	7.41	51.3	5.40	25.0	7.44	34.3	12.6	42.7	17.1	28.2	2.93	60.4	7.01	39.5	
7	11.9	32.0	13.4	34.4	12.1	31.8	11.3	10.9	9.44	25.6	14.5	32.4	16.5	26.7	7.69	54.2	10.4	26.8	
1	12.4	55.4	12.7	46.2	10.3	19.4	10.2	17.6	9.86	21.2	15.4	40.7	22.8	35.8	4.82	51.4	8.01	16.6	0
3	15.7	43.4	17.6	30.0	12.9	25.5	14.0	10.3	12.7	20.7	21.6	38.4	23.6	39.5	8.59	53.2	13.2	26.5	
20	16.1	43.0	16.3	31.4	13.8	21.8	14.0	9.3	13.6	29.9	14.4	25.2	25.8	32.8	9.20	45.6	12.8	34.8	10
16	32.3	22.5	34.4	18.2	31.3	8.7	32.4	12.9	28.2	14.9	28.1	20.6	35.2	23.3	26.7	47.1	32.0	15.9	30
12	48.0	15.4	49.3	10.4	49.6	10.7	51.1	7.5	60.2	33.2	46.6	12.7	40.8	15.6	47.8	16.7	49.5	11.8	
15	59.6	16.7	60.1	14.0	60.5	3.9	63.0	5.3	55.9	15.6	50.7	6.7	52.0	20.0	54.6	22.2	61.8	10.3	60
17	70.0	15.4	70.4	11.1	72.4	9.2	74.7	2.9	63.8	15.5	58.0	7.7	75.1	21.6	52.5	56.3	76.7	18.3	
5	70.5	13.4	72.3	9.4	73.0	8.2	73.2	2.5	66.0	9.7	61.9	11.4	67.6	9.8	70.2	23.4	73.1	7.5	
4	70.8	12.8	73.0	10.6	75.1	4.0	75.7	6.8	65.5	14.3	61.8	5.2	69.7	16.7	68.4	14.7	75.6	7.5	
2	73.8	12.0	75.6	11.1	74.5	4.3	77.6	6.2	71.3	14.6	70.5	6.7	68.6	17.8	68.2	13.8	79.8	8.2	
11	119	14.2	119	12.5	124	10.1	124	3.3	115	17.6	112	7.1	107	12.1	113	30.3	132	8.9	120
19	128	13.3	121	14.9	124	4.0	132	3.7	154	13.6	129	7.0	117	14.8	129	16.4	138	14.0	
8	151	12.3	151	14.9	153	12.4	159	5.5	152	12.6	145	5.5	142	10.7	117	50.7	167	10.4	
9	278	12.3	271	9.2	269	24.2	295	10.0	248	(n = 1)	267	8.0	282	24.8	283	6.4	293	4.4	300

Tab. 8. Mean coefficients of variation – Precision prifile – T_4-EQCS.

Range (nmol · liter^{-1})	All n = 103	Own method n = 18	Kit 1 n = 6	Kit 2 n = 10	Kit 3 n = 5	Kit 4 n = 8	Kit 6 n = 9	Kit 11 n = 8	Kit 12 n = 7
mean c.v.									
(0–300)	34.8	27.6	22.8	13.0	22.9	20.7	24.6	41.1	22.4
0–40	54.3	43.3	36.0	20.7	28.8	33.6	32.8	57.1	34.6
40–120	14.3	11.3	7.38	5.20	17.2	8.21	16.0	28.5	10.6
> 120	12.6	13.0	12.7	5.63	13.1	6.83	17.8	11.4	9.42

This table, as in table 6, shows precision of each method in different parts of the standard curve.

Tab. 9. Analysis of regression data from T_3- and T_4-EQCS. Regression data for the equation $y = a + bx$, the first term of the two parameters is entered as x.

Limits	Values from own standard curve compared with those from Curve I		Values read off Curve I compared with those read off Curve II	
	T_3 (n = 91)	T_4 (n = 101)	T_3 (n = 91)	T_4 (n = 101)
Correlation coefficient				
r > 0.99	80 (88)	90 (89)	86 (95)	91 (90)
r 0.95–0.99	7 (8)	7 (7)	5 (5)	6 (6)
r < 0.95	4 (4)	4 (4)	0	4 (4)
Slope of the regression line				
b > 1.10	10 (11)	5 (5)	16 (17)	49 (49)
b 0.90–1.10	74 (81)	87 (86)	49 (54)	38 (38)
b < 0.90	7 (8)	9 (9)	26 (29)	14 (13)
y-axis intercept				
T_3 / T_4				
a > 0.2 / a > 10	3 (3)	1 (1)	22 (24)	3 (3)
a ± 0.2 / a ± 10	78 (86)	99 (98)	43 (47)	69 (68)
a < −0.2 / a < −10	10 (11)	1 (1)	26 (29)	29 (29)

The above data exclude 18 T_3- and 23 T_4-EQCS participants who returned insufficient data. The figures in brackets represent the percentages of the useable data.

Tab. 10. Comparison of binding of serum no. 1 with own-method zero standard. All figures are expressed as percentage (Own-method/Serum no. 1).

Range	T_3-EQCS	T_4-EQCS
> 110%	4 (4)	0
90–110%	68 (76)	59 (55)
< 90%	18 (20)	48 (45)

Data from 90 T_3-Curves and 107 T_4-Curves. Figures in brackets indicate percentage of processed data.

Tab. 11. Superimposibility of Curve I upon Curve II, a subjettive comparison.

Superimposition	T_3-EQCS	T_4-EQCS
Good	34 (38)	34 (32)
Average	50 (56)	60 (56)
Bad	6 (6)	13 (12)

Data from 90 T_3- and 107 T_4-standard curves. Figures in brackets denote percentage of processed data.

tion of the cases passed near to the origin (T_3 = 86%, T_4 = 98%).

The corresponding data for the comparison between curve I and curve II are as perhaps expected not so good. The correlation between the two curves was similar to that between the own curve and curve I (T_3 = 95%, T_4 = 90%) although both the slope of the regression line varied widely, with only 54% of T_3 and 38% of T_4-EQCS participants lying between the limits b = 0.9–1.1. The number of regression lines passing through or near the origin was also fewer (T_3 = 47%, T_4 = 68%).

Comparison of Binding of Tracer to the Kit — Own Method and Hidden Standard Curve Zero Standards

Table 10 shows the comparison of the binding of the tracer to the own-method zero standard with the binding to the hidden standard curve zero standard. The ideal value is 100%. The range chosen as acceptable in both cases was between 90 and 110%. 76% of the T_3-EQCS participants and 55% of those taking part in the T_4-EQCS lay within this range.

Comparison of Curve I and Curve II for Superimposibility

Table 11 shows a subjective comparison of the kit/own-method and hidden standard curves with regard to their superimposibility. Superimposibility was seen in 38% of the T_3-standard curves and in 32% of the T_4-standard curves. The division into good, average and poor superimposibility was purely subjective, no detailed mathematical analysis being performed.

Discussion

As in the previous EQCS, the majority of participants used commercial kits, with or without modification, for the determination of total T_3 and total T_4 in serum. At the time of writing, no less than 30 kits are commercially available for these two assays. The performance of the kits, although not as bad as those 5 years ago (2), varies widely as seen in the precision profiles in table 6. Although many kits have standard curves in human serum, there are still several which use either animal serum or even buffer as the matrix for the standards. A typical result of a potentially good kit spoilt by having standards in buffer is the case of kit 10 in the T_3-EQCS. As seen in table 6, the precision is good, but when one looks at table 5, the accuracy is lacking. That the kit components and methodology are otherwise in order can be seen from the comparison of the results read off curve I (fig. 1a) and those read off the hidden standard serum curve (curve II) as seen in figure 1b. The results, when read off the serum standard curve are not only precise, but accurate. Only one kit-producer (kit 2) had a product which was precise and accurate in the regions of interest in both T_3 and T_4 assays. This kit was robust and also one of the cheapest on the market, although it used a double-antibody method of separation of bound and free antigen.

The EMIT assay for T_4 showed precision and accuracy comparable with some RIA methods, although it is difficult to draw too many conclusions from such a small number of data.

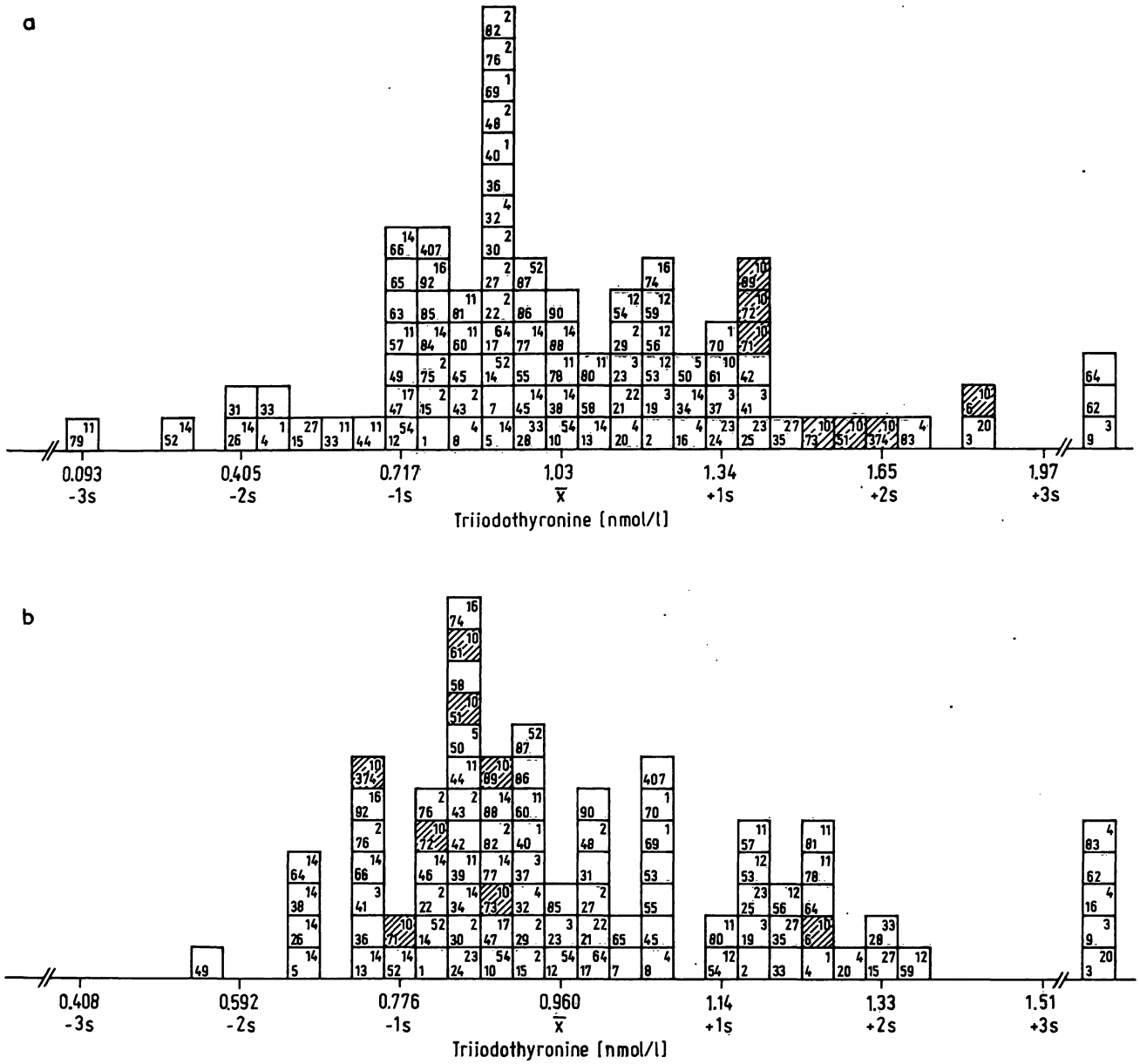

Fig. 1. Histogrammes from the T_3-EQCS showing results from Kit 10 (shaded squares).
a) read off curve I (from participant's own data worked out using spline function) It is plain to see that the buffer standards in the kit give biased results.
b) Read off curve II, the hidden standard curve in human T_3-free serum. With the exception of one result, the values obtained for serum No. 4 fall in the main block of results, thus demonstrating that the standards in the kit are responsible for the positive bias seen in fig. 1a.
The figures in the boxes are: lower left, participant number and upper right, kit number.

The performance of different types of laboratory dispelled the myth that only certain types of laboratory can measure T_3 and T_4. This is perhaps to be expected, as the majority of participants use kits, where it is only necessary to follow the recipe and work with precision. If this survey represents a true cross-section of the laboratories performing T_3 and T_4 assays in the Federal Republic of Germany, then the general RIA laboratories performing these tests outnumber the nuclear medicine/radiology laboratories by about 3:1.

The regression analysis and related data (tables 9–11) demonstrate that most laboratories can work out their data efficiently. Those having the worst correlations were almost inevitably those laboratories who used either a computer or desk-top calculator to evaluate data without a visual display of the standard curve.

The results emphasise the need for a visual check of what the calculator is doing.

The distribution of the values for r, a and b for the regression data for the T_3-EQCS was *Gaussian*, whereas those for the T_4-EQCS were skewed. The latter can probably be best explained that the T_4-free serum used in the hidden standard curve was not really T_4-free, but

contained about 10 nmol · liter^{-1} T_4 – although this was checked in no less than 3 systems before lyophilisation. One thing is evident from this EQCS, and that is, all T_3-kits are able to measure an elevated T_3-level in serum and thus detect a case of T_3-hyperthyroidism. Concomitantly, most T_4-kits are able to differentiate between euthyroid and hypothyroid serum levels, although only one kit is able to differentiate between no T_4-output and minimal T_4-output from the thyroid gland, an important point in deciding whether the thyroid is absent or whether cells are still present which are producing thyroid hormones, an important fact in the follow-up of cases of malignant thyroid carcinoma after total thyroidectomy.

Inspite of these findings, there is still room for improvement in the design of many T_3 and T_4-kits and methods, especially in their "robustness". As *Ekins* has so often said, a kit must be able to perform well, that is, precisely and accurately under all possible conditions.

References

1. Marschner, I., Bottermann, P., Erhardt, F., Linke, R., Löffler, G., Maier, V., Schwandt, P., Vogt, W. & Scriba, P. C. (1974), Horm. Met. Res. *6*, 293–296.
2. Horn, K., Marschner, I. & Scriba, P. C. (1976), J. Clin. Chem. Clin. Biochem. *14*, 353–360.
3. Marschner, I., Erhardt, F. W. & Scriba, P. C. (1976), J. Clin. Chem. Clin. Biochem. *14*, 345–351.
4. Wood, W. G., Bauer, M., Marschner, I. & Scriba, P. C. (1980), J. Clin. Chem. Clin. Biochem. *18*, 183–192.
5. Ekins, R. (1979), Assay Design and Quality Control: In Radioimmunology 1979 (Ed. Ch. A. Bizollon), Elsevier North Holland, pp. 239–255.
6. Scriba, P. C., Bauer, M., Emmert, D., Fateh-Moghadam, A., Hofmann, G. G., Horn, K. & Pickardt, C. R. (1979), Acta Endocrinol. (Kbh.) *91*, 629–643.

Dr. W. G. Wood, Ph. D.
Klinik für Innere Medizin
Medizinische Hochschule Lübeck
Ratzeburger Allee 160
D-2400 Lübeck 1